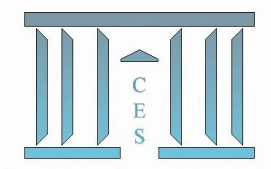
The Spectral Stress VaR (SSVaR)

Dominique Guegan, Bertrand K. Hassani, Kehan Li

UNIVERSITÉ PARIS 1
PANTHÉON SORBONNE

Documents de Travail du
Centre d'Economie de la Sorbonne

CES Working Papers



The Spectral Stress VaR (SSVaR)

Dominique GUEGAN, Bertrand HASSANI, Kehan LI

2015.52

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
http://centredeconomiesorbonne.univ-paris1.fr/
ISSN : 1955-611X

The Spectral Stress VaR (SSVaR)

D. GUÉGAN*[1], B. HASSANI † ‡[2], and K. LI §[1]

[1]Université Paris 1 Panthéon-Sorbonne, CES UMR 8174, 106 boulevard de l'Hopital 75647 Paris Cedex 13, France.

[2]Grupo Santander and Université Paris 1 Panthéon-Sorbonne CES UMR 8174, 106 boulevard de l'Hopital 75647 Paris Cedex 13, France.

Abstract

One of the key lessons of the crisis which began in 2007 has been the need to strengthen the risk coverage of the capital framework. In response, the Basel Committee in July 2009 completed a number of critical reforms to the Basel II framework which will raise capital requirements for the trading book and complex securitisation exposures, a major source of losses for many international active banks. One of the reforms is to introduce a stressed value-at-risk (VaR) capital requirement based on a continuous 12-month period of significant financial stress (Basel III (2011) [1]. However the Basel framework does not specify a model to calculate the stressed VaR and leaves it up to the banks to develop an appropriate internal model to capture material risks they face. Consequently we propose a forward stress risk measure "spectral stress VaR" (SSVaR) as an implementation model of stressed VaR, by exploiting the asymptotic normality property of the distribution of estimator of VaR_p[1]. In particular to allow SSVaR incorporating the tail structure information we perform the spectral analysis to build it. Using a data set composed of operational risk factors we fit a panel of distributions to construct the SSVaR in order to stress it. Additionally we show how the SSVaR can be an indicator regarding the inner model robustness for the bank.

Keywords: Value at Risk, Asymptotic theory, Distribution, Spectral analysis, Stress, Risk measure, Regulation.

*dguegan@univ-paris1.fr

†Disclaimer: The opinions, ideas and approaches expressed or presented are those of the authors and do not necessarily reflect Santander's position. As a result, Santander cannot be held responsible for them.

‡bertrand.hassani@gmail.com

§kehanleex@gmail.com

[1]The VaR_p called Value at risk is the classical quantile risk measure computed from a probability distribution for a given confidence level $0 < p < 1$.

Documents de travail du Centre d'Economie de la Sorbonne - 2015.52

1 Introduction

Efforts are underway by national and supra-national supervisors such as the EBA as well as international organizations involved in financial stability - the International Monetary Fund (IMF), the Basel Committee for Banking Supervision (BCBS) or the European Central Bank (ECB) - to develop new risk-modelling techniques and stress test methodologies to better identify and capture the risks that trigger economic and financial instability. The Basel Committee's reforms in Basel III (2011) [1] have the objectives to strengthen global capital and liquidity rules with the goal of promoting a more resilient banking industry. One of the reforms is to propose a stress risk measure as an internal model for the bank, that is a "stressed VaR" based on a continuous 12-month period of significant financial stress. Until now it seems that few articles focus on the concept of the stressed VaR: Alexander and Ledermann (2012) [4] define the stressed VaR through the VaR computed on a stressed data sample from historical data set tranformed using a stressed covariance matrix; Abdymomunov, Blei and Ergashev (2014) [2] stress the loss distribution of a data set selecting a stress scenario on the historical data set and compute the corresponding VaR. Guégan and Hassani (2015) [3] integrate the stress by fitting a normal-inverse Gaussian (NIG) distribution instead of a Gaussian distribution on the returns and then derive the VaR for market risk. Alternative scenarii have been employed: Hoggarth, Sorensen and Zicchino (2005) [12] stress the coefficients of a vector autoregression model (VAR) using specific shocks; Fong and Wong (2008) [7] combine two Gaussian VAR models capturing the information in the tails, introducing also a form of stress.

In this paper, we are interested in investigating the probabilistic and statistical properties of the r.v. $\widehat{VaR}_p$ [2], computed from the estimated distribution associated to the risk factor, and in building the associated confidence interval around the true VaR in order to propose a stress risk measure. In the following this stress risk measure is referred to the Spectral Stress Value at Risk (SSVaR). We show that this SSVaR provides a robust prudential risk measure which takes into account a stress environment which allows to take the information in the tails with different intensities thanks to the spectral analysis and which provides risk managers and regulators with an alert indicator encompassing the randomness and the stress from the VaR_p.

To achieve these objectives, we first recall the distribution of the r.v. $\widehat{VaR}_p$ and second we discuss the implementation of these assumptions on real data sets. Then we illustrate how we can build the SSVaR with real data and how it can be used as an alert indicator for risk management.

This article is structured as follows. In Section two the asymptotic normality property of $\widehat{VaR}_p$ is presented, from which we develop the associated confidence interval.

[2] It is defined in the Theorem 2.1 in section 2.

We discuss the theoretical result provided by the Theorem 2.1 using simulations. Then the SSVaR indicator is defined. In the third section we present the data set we use to illustrate the building of our indicator. We analyse several aspects for risk management based on the goodness-of-fit of different distributions, the prudential risk measure requirement, the construction of the confidence interval of VaR_p and its convergence. We build the SSVaR indicator and show how it works. Section four concludes. In some annexes we have postponed the proof of the Theorem and we recall some of the techniques we use for the simulation and estimation strategies.

2 Methodology

In this section we present the theoretical foundations of our methodology.

2.1 Theoretical Foundation

Considering a set of i.i.d. r.v. characterized by a distribution law F, Rao (2001) [8] provides the distribution law of the r.v. $\widehat{VaR}_p$ computed from this set of r.v. for a given p. We recall this theorem.

Theorem. *Let $X_1, ..., X_n$ be a sequence of i.i.d r.v. whose continuous and strictly monotonic distribution is F, f the associated continuous density, $0 < p < 1$ a given number and np is assumed not to be an integer, if we denote ξ_p the quantile associated to F at level p, then the order statistic $\hat{\xi}_{p_n}$ has the following property:* [3]

$$\hat{u}_{p_n} = \frac{\hat{\xi}_{p_n} - \xi_p}{\sqrt{V}} \xrightarrow{distribution} N(0,1), \quad as \quad n \longrightarrow \infty \tag{2.1}$$

where

$$V = \frac{p(1-p)}{f(\xi_p)^2 n}. \tag{2.2}$$

For the proof we refer to Rao (2001) [8]. This theorem tells that the asymptotic distribution of the r.v. $\widehat{VaR}_p$ is Gaussian, and the convergence speed depends on the parameter p, the density f and the sample size n. The following corollary provides the speed of convergence:

Corollary. *Under the same assumptions given in the previous theorem, assuming that F has a bounded second derivative F'', then:*

$$\sup_{t \in \mathbb{R}} |P\left(\frac{\hat{\xi}_{p_n} - \xi_p}{\sqrt{V}} \leq t\right) - \Phi(t)| = O(n^{-\frac{1}{2}}), \quad as \quad n \longrightarrow \infty \tag{2.3}$$

where V is given in (2.2), and Φ is the cdf of standard Gaussian distribution.

[3] The quantile ξ_p corresponds to the VaR_p, thus the distribution of $\widehat{VaR}_p$ based on the order statistics of $X_1, ..., X_n$ is provided by this theorem. Note that $\hat{\xi}_{p_n} := X_{([np]+1)}$.

This corollary says that the distribution of the r.v. $\hat{\xi}_{p_n}$ converges to the Gaussian distribution with the speed $n^{-\frac{1}{2}}$. Then for any given $0 < q < 1$ [4], we have $P(|\hat{u}_{p_n}| \leq z_{\frac{1+q}{2}}) \approx q$ for n large enough, and $z_{\frac{1+q}{2}}$ associated with $\frac{1+q}{2}$ is provided by the table of the standard Gaussian distribution.

In the following we are interested to use the result of Theorem 2.1 to build a confidence interval around ξ_p. In order to verify the robustness of this building, we analyse the influence of the parameters used to compute the variance V introduced in (2.2).

2.2 Convergence of the distribution of $\widehat{VaR}_p$ to the Gaussian distribution

To verify the influence of the parameters f, n and p in the asymptotic convergence of $\widehat{VaR}_p$ towards the true VaR_p, we propose a plan of simulations making varying these parameters which appear in the computation of the variance V in (2.2). Thus, we carry out the following experiment. [5]

1. Given a density f, $0 < p < 1$ and n a sample size, we generate $1000n$ samples from f and obtain a $n \times 1000$ matrix.

2. We sort the elements of each column of this matrix in ascending order. We consider the $([np] + 1)^{th}$ row of this matrix composed by 1000 realizations of $\widehat{VaR}_p$: we call this row $\Psi = \left(\widehat{VaR}_p^i\right)_{i=1,\ldots,1000}$.

3. We denote $\bar{\Psi} = \frac{1}{1000} \sum_{i=1}^{1000} \widehat{VaR}_p^i$, the mean of the row.

4. In equation (2.2), we replace ξ_p by $\bar{\Psi}$ and we still note V the corresponding variance.

5. We center and normalize each term of this row in the following way: $\left(\frac{\widehat{VaR}_p^i - \bar{\Psi}}{\sqrt{V}}\right)_{i=1,\ldots,1000}$.

6. We use a QQ plot to compare these new 1000 quantiles with the 1000 quantiles of a Gaussian law, mean 0, variance 1. We test the fit using the Kolmogorov-Smirnov test.

[4] As we fix p, $\hat{\xi}_{p_n}$ is a r.v. with the asymptotic normality property. Thus we use this property to build a confidence interval around ξ_p with a confidence level q, $0 < q < 1$.

[5] Another approach would be the use of the Δ method, Gao and Zhao (2011) [11] It will be the purpose of another paper.

We do these six steps for five densities f, for n varying from 501 to 1000001, and for $p = 0.95, 0.999$. We use densities which are asymmetric, leptokurtic and fat-tail: the lognormal distribution ($\mu = 7.9491$ and $\sigma = 1.2373$), the Generalized hyperbolic distribution (GH) ($\alpha = 2.8367e-5$, $\beta = 2.7881e-5$, $\delta = 1000$, $\mu = 10000$ and $\lambda = -0.5$), the Generalized Pareto distribution (GPD)($\xi = 1.0080e-2$, $\beta = 3.2245e+5$ and $\mu = 314391.3$), the Extreme value distribution (GEV) ($\xi = 0.9853$, $\beta = 2635.2984$ and $\mu = 2583.3371$) and the $\alpha - stable$ distributions ($\alpha = 0.7720$, $\beta = 0.95$, $\gamma = 908.5263$ and $\delta = 1655.948$). Results are provided in figure 1 [6].

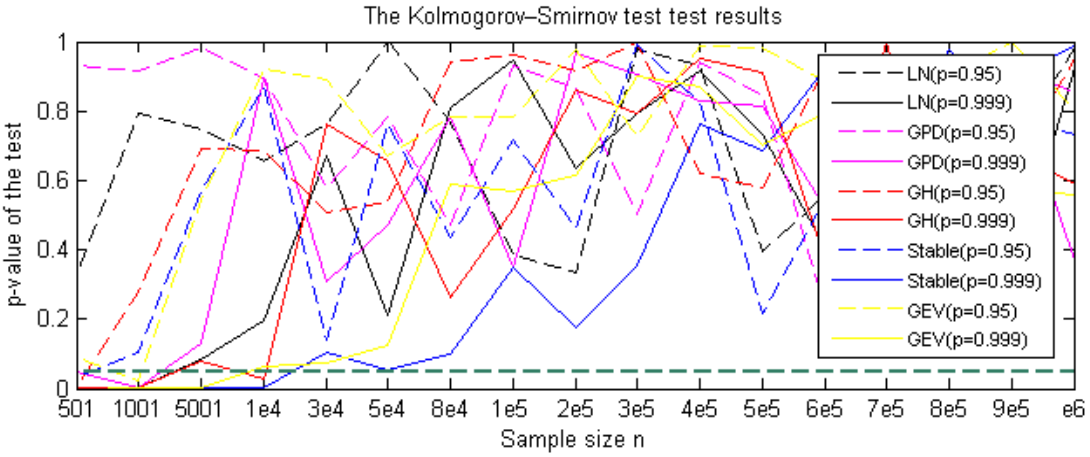


Figure 1: P-value of the K-S test run for different values of n [8], for five densities (lognormal, GH, GPD, GEV and $\alpha - stable$ and for $p = 0.95, 0.999$.

We observe that, as soon as, $n > 30001$, whatever the densities, and whatever p, the convergence to the Gaussian distribution works. For $5001 < n < 30001$ there is no convergence for the GEV, GH and $\alpha - stable$ distributions when $p = 0.999$. For $1001 < n < 5001$, there is no convergence for the GEV, GH, lognormal and GPD distributions when $p = 0.999$. For $501 < n < 1001$, we have only convergence for the lognormal and the GPD distributions when $p = 0.95$. We observe that, for finite samples (less that 1001) the convergence is never attained for $p > 0.95$. When the convergence works it is only for the lognormal distribution and the GPD (which is known to be very unstable) when $p \leq 0.95$. Now as soon as we use a fat-tail distribution with asymmetry behaviour the convergence is attained only for a large n.

These results show that the Theorem 2.1 can fail as soon as we work with data sets characterised by distribution with fat-tail and also when we are interested to analyse the behaviour of the quantiles in the tail (p large). Nevertheless, in this paper

[6] Other figures can be provided under request.

[8] n=501,1001,5001, 10001,30001,50001,80001,100001,200001,300001,400001,500001,600001,700001, 800001,900001,1000001.

we will use the result of the Theorem 2.1 to propose an alert indicator based on a confidence interval built around $\widehat{VaR}_p$. This is the purpose of the next paragraph.

2.3 An alert indicator called SSVaR

From the Theorem 2.1, we can build a confidence interval CI_q with $0 < q < 1$ around the true unknown VaR_p:

$$\boldsymbol{\xi_p} \in \left[\hat{\boldsymbol{\xi}}_{\boldsymbol{p_n}} - \boldsymbol{z}_{\frac{1+q}{2}}\sqrt{\boldsymbol{V}}, \quad \hat{\boldsymbol{\xi}}_{\boldsymbol{p_n}} + \boldsymbol{z}_{\frac{1+q}{2}}\sqrt{\boldsymbol{V}}\right] \tag{2.4}$$

We are going to use this confidence interval to define an alert indicator permitting to the managers to know when a sequence of events are “anormal" in the sense that the measure associated to the corresponding risk factor is outside a “certain" confidence level.

For a sequence of p, $p_1 < p_2 < ... < p_k$, we get a sequence of $VaR_{p_i}, i = 1, ..., k$. For each VaR_{p_i}, we can build around this value a confidence interval CI_{q_i,p_i}, for a given q_i, $i = 1, ..., k$. The parameters q_i and p_i can be equal or different. Now, we consider the area between each VaR_{p_i} and the upper bound of its corresponding CI_{q_i,p_i}: we illustrate this area in Figure 2.

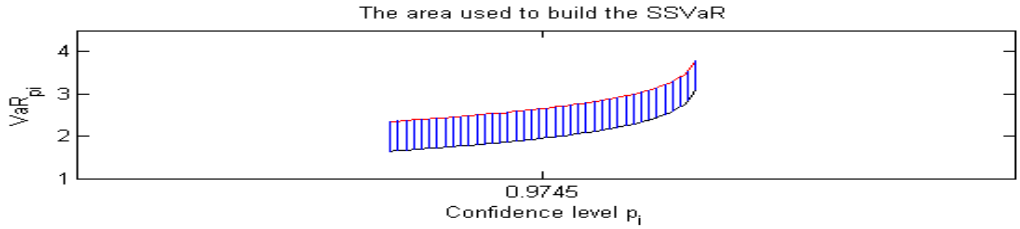


Figure 2: We provide the area corresponding to the SSVaR, the lower bound corresponds to VaR_{p_i} and the upper bound to the upper bound of $CI_{q_i,p_i}, i = 1, ..., k$.

This area - delimitated between VaR_{p_i} and the upper bound of CI_{q_i} - corresponds to the Spectral Stress VaR measure we propose to use as an alert indicator. Indeed, having the VaR for different p provides us with the spectral VaR (SVaR). The construction of a set of confidence intervals around the SVaR provides us with an acceptable range of variation for the VaR_p-s. Only considering the upper bounds on the confidence intervals gives the spectral stress VaR or $(SSVaR_{p,q})$. The upper envelop allows us to use multiple values at a particular percentile level.

3 Application

In this section, the methodology presented earlier is applied to real data. The first step is to fit various distributions to the data. In a second step, the Confidence intervals around the VaRs are built. Finally, in a last step, the SSVaR is created.

3.1 Data sets and fit

In order to build the confidence interval around the VaRs, two data sets Ω_1 and Ω_2 based on a sample of operational risk losses [9] whose lengths are respectively $n = 3109$ $(01/01/2013 - 29/11/2013)$ and $n = 32727$ $(24/01/2007 - 31/12/2013)$ are used as common basis in this exercise. Table 1 provides the descriptive statistics of these two samples.

	mean	std	skewness	kurtosis
Ω_1	9.5006e+3	4.4428e+4	12.5054	179.5641
Ω_2	8.6227e+3	4.2775e+4	21.3599	754.5112

Table 1: First four empirical moments for Ω_1 and Ω_2.

From table 1 we observe that both Ω_1 and Ω_2 are characterised by a positive skewness and a large kurtosis (larger than 150). It means that the data are asymmetric and leptokurtic and therefore the distributions used to characterise them should be selected such that they are consistent with these properties. We will use six distributions including the lognormal fitting (as it is widely used by practitioners) and the non-parametric adjustment (as it is the most representative of the data).

The non parametric adjustment is obtained using a cross validation strategy ([10], [5] and [9]) considering a Gaussian Kernel (Appendix B and C). Regarding the parametric fittings, five distributions are considered: the lognormal, the GH, the GPD, the GEV and the $\alpha - stable$ distributions. The parameters of the lognormal, GH, GPD (except the location) and GEV distribution are estimated by maximum likelihood. The location parameter of the GPD is estimated using a Hill plot, while for the parameters of the $\alpha - stable$ distribution the McCulloch method [6] has been implemented.

The standard deviation of the estimated parameters for the lognormal and the GEV distributions are computed using the inverse of the Hessian matrix. For GH, GPD and $\alpha - stable$ distributions, the standard deviations are computed by bootstrapping. The estimated parameters of these distributions are presented in table 3 (for Ω_1)

[9] Provided by an Eastern European financial institution

and table 4 (for Ω_2) [10]. According to the K-S tests, no distribution has been found acceptable at the 5% level.

Remark 3.1. *Operational Risk Data are usually multimodal and consequently it is generally quite difficult to properly fit an unimodal distribution to them, nevertheless in the following we will work with the distributions aforementioned as this does not affect the construction of the SSVaR later on, and the results can be extrapolated to any other kinds of distributions.*

As we see in Section 2.2 that the theorem 2.1 we use to build the confidence interval around VaR_p is not always true depending mainly on the distribution we use, we verify with our data set when this convergence exists using the six previous distributions and making n varying from 501 to 1000001. We consider also different values for p and q from 0.95 to 0.999.

Our computations show that the convergence is observed as soon as $n = 30001$. This confirms the results obtained by simulations in the previous section. It is the reason for which we use two data sets in this exercise, one with finite distance ($n = 3109$) and another one with $n = 32727$, which is close to the asymptotic setting.

3.2 Confidence interval around $\mathbf{VaR}_p$

Given the six previous fitted distributions $f_i, i = 1, ..., 6$, we build for each distribution a sequence of confidence interval CI_q around VaR_p using equation (2.4), for the two data sets Ω_1 and Ω_2. This means that we make varying n, taking $n = 3109$ for Ω_1 and $n = 32727$ for Ω_2.

In practice, to build CI_q we proceed in the following way: for each $f_i, i = 1, ..., 6$, using the algorithm 2.2, steps 1 to 4, we build the sequence $\left(\widehat{VaR}_p^i\right)_{i=1,...,1000}$. For fix p, we build the confidence interval CI_q, around VaR_p, given by (2.4) replacing $\hat{\xi}_{p_n}$ by $\bar{\Psi}$ and using V computed in steps 3 and 4. Then, we get a confidence interval at level q around $\widehat{VaR}_p$.

We can compute, using the sequence $\left(\widehat{VaR}_p^i\right)_{i=1,...,1000}$ how many times $\widehat{VaR}_p^i$ are outside this confidence interval: we denote this number m. In Table 5 ($n = 3109$) and Table 6 ($n = 32727$) we provide for each distribution and different values of p on each line - the value of $\bar{\Psi}$, the CI_q and in bracket $\frac{m}{1000} = \hat{q}$. This value of $\hat{q}$ corresponds to the percentage of events outside the confidence interval that we have

[10] Fit results for the other subsets of Ω_2 are available upon request.

to compare with $1-q$.

For instance in table 5 ($n = 3109$), using the lognormal, for $p = q = 0.95$, we get $\hat{q} = 0.041$. In practice we expect to have less than $1 - q = 5\%$ of losses outside the confidence interval. In that case no alert arrives. For this distribution the result is the same when $n = 32727$ (Table 6). Now if we fit a GH distribution, for $p = q = 0.95$, $\hat{q} = 0.087$ in Table 5 and $\hat{q} = 0.055$ in Table 6. Thus, with this fitting we will have an alert whatever the value of n.

Remark 3.2. *Another way of analysing the behaviour of the confidence interval applied to real data is to rely on the number of data points used to fit the distributions (n_Ω) and not anymore on the number of points randomly generated (n). Indeed, the reliability of the VaR evaluated for a risk factor would not be biased by a large random generation and would therefore integrate the robustness of the distributions fitted to these data - the lower the number of data point, the less reliable the VaR.*

We provide now some comments concerning the result obtained for the confidence interval around VaR_p for each distribution fitted. Indeed, in Tables 5 and 6 we observe that:

1. the lognormal fit tends to underestimate the VaR compared to empirical quantiles, while Kernel smoothing and GPD tend to provide a much higher VaR. Though the rank of the VaR - with respect to the distributions used to compute it - is quite unstable. The presence of extreme incidents in Ω_2 explains the large computed VaR.

2. The length of CI_q depends on the distribution. Fat tailed distribution such as the GPD tends to lead to larger confidence interval but as exhibited in table 6, this is not always the case as with the 99.9th percentile; in another hand the $\alpha - stable$ distribution induces a larger interval.

3. Concerning the convergence of $\hat{q}$ to $1 - q$:

   (a) From table 5 (set Ω_1) we observe that the $\hat{q}$ does not converge at all for the GH fit, while it only converges for the GPD, GEV and $\alpha - stable$ fits when $p = q = 0.999$.

   (b) From table 6 (set Ω_2) we observe that the $\hat{q}$ always converge to $1 - q$.

3.3 The SSVaR alert indicator in practice

In Figure 3 we illustrate how we can build the SSVaR indicator.

First, we illustrate some findings we already discuss: in the top, left part of Figure 3 we represent for each distribution fitted in Ω_2 (non-parametric, lognormal, GPD, GH and $\alpha - stable$), the evolution of the $\hat{VaR}_p$ quantile for $0.95 \leq p \leq 0.999$ [11]. Definitely the VaR_p is always larger with GPD, and the smallest with the lognormal distribution. Nevertheless, we observe that when $p = 0.999$ the largest VaR is obtained with the $\alpha - stable$ distribution. For the other values of p ($0.95 < p < 0.999$) the ranking of the distributions providing the smallest VaR_p to the largest VaR_p is: the lognormal, then the empirical, then the GH, then the $\alpha - stable$, then the Kernel and then the GPD.

Now concerning the SSVaR indicator, from Figure 3 we obtain several kinds of information.

1. On the bottom part, left part, we have built the envelop obtained when we fit the lognormal distribution using the Ω_2 set: thus for $0.95 \leq p \leq 0.999$, we represent as the lower VaR_p, and the upper bound is the upper bound of the CI_q (built with $q = p$ each time). The envelop corresponds to the black picture.

2. On the bottom part, right part, we have built the envelop using the GPD distribution fitted with the Ω_2 set. The building is the same as in 1).

3. Looking at these two graphs we observe several behaviours: Using the GPD distribution the upper bound is higher than using the lognormal distribution (be careful of the scale), and the width of the envelop is larger with the GPD than with the lognormal distribution, whatever p.

4. In Figure 3, on the top right part, we provide in more details the evolution of the width of the envelop for each distribution making varying p. The largest width is obtained with GPD, the smallest with lognormal. We observe also that for the other distributions we have no uniform behaviour. For instance using the Kernel the width is nearly constant. The width of the envelop with $\alpha - stable$ distribution is always larger than the width obtained with the GH distribution.

These different remarks show the great influence of the choice of the distributions used to characterise the features of the risk factor in order to build a correct strategy for risk management.

Now, we propose a way to use the SSVaR to provide risk managers and regulators an alert indicator which encompasses the randomness and stress from the VaR_p.

[11] For p varies, we compute the $\widehat{VaR}_p$ by definition using directly Ω_2. It's empirical benchmark.

The idea is to use the SSVaR to capture the probability of the VaR to be inaccurate and to provide an acceptable range of values to stress it.

Considering the data set Ω (length n_Ω), we begin with a specific subset of Ω which will be our basic information set to create the original envelop SSVaR. We call it Ω_0. From Ω_0, we fit a distribution F with density f. Then, for a given sequence of p and q, we build the envelop delimitated by VaR_p and the upper bound of each $CI_{p,q}$. Then, we increase the data set Ω_0 of 3 months for instance, and obtain a new data set Ω_1. We do exactly the same job as before and we verify if the new envelop is inside the previous one. If it is inside, no alert. If it is outside, we have alert and so on.

We provide now an example built with sub-samples coming from the set Ω_2, on which we fit $\alpha - stable$ distributions. The first set we use is Ω_0 corresponding to data from January 4th, 2011 to December 31th, 2011 ($n = 2572$). Using this set we compute the SSVaR for $p = 0.95, 0.975, 0.99, 0.995$ (here $p = q$).

Then, we add to this set 9 months and get the set Ω_1. We use data from January 4, 2011 to September 30, 2012 ($n = 6195$). Using this data set we build the SSVaR for the same p and q as before. We do the same job for a new set Ω_2 corresponding to the period January 4th, 2011 to December 12th, 2012 ($n = 7189$).

The results are provided in Table 2 lines 1 to 3 and illustrated in Figure 4. Considering the values obtained to delimitate the bounds of SSVaR and comparing these values to the benchmark provided by Ω_0, we observe that all the SSVaR computed from Ω_1 and Ω_2, whatever p, are always smallest than the values which define the SSVaR computed with Ω_0. So with these data sets, we do not trigger an alert.

Now we propose a stress approach, building a data set Ω_0^{stress} such that we add to the set Ω_0 the sequence of points corresponding to the period January 24th, 2007 to September 30th, 2007 ($n = 4851$). we compute the SSVaR for this set and we observe that as soon as $p > 0.95$, the bounds of these SSVaR are outside the bounds of the SSVaR computed using Ω_0.

We do again the same job, building another stress set $\Omega_0'^{stress}$, using still Ω_0 and adding the period January 24th, 2007 to December 31th, 2007 ($n = 5451$). We compute the SSVaR and again, the bounds of these SSVaR are bigger than these computed using Ω_0. The results are provided in Table 2 lines 4 and 5, and illustrated in Figure 4.

Thus, for these two data sets Ω_0^{stress} and $\Omega_0'^{stress}$, we will have an alert corresponding to the arrival of extreme events, when we fit an $\alpha - stable$ distributions on these data sets.

We have done the same job using a lognormal distribution with the same stress sets and we have observed that we never get an alert.

	p=0.95	p=0.975	p=0.99	p=0.995
Ω_0	[3.8, 4.0]	[8.9, 9.8]	[29.4, 34.9]	[73.5, 95.0]
Ω_1	[3.5, 3.7]	[8.0, 8.8]	[25.8, 30.6]	[65.1, 84.2]
Ω_2	[3.4, 3.6]	[7.7, 8.5]	[25.0, 29.6]	[62.1, 80.1]
Ω_0^{stress}	[3.7, 3.9]	[9.0, 9.9]	[30.3, 36.1]	[78.9, 102.9]
$\Omega_0'^{stress}$	[3.7, 3.9]	[9.0, 9.9]	[31.0, 36.9]	[80.6, 105.2]

Table 2: SSVaR for 5 subsets of Ω_2 on which the $\alpha - stable$ has been fitted and for 4 different values of p. Ω_0 corresponds to the period January 4th, 2011 to December 31th, 2011 ($n = 2572$); Ω_1 corresponds to the period January 4, 2011 to September 30, 2012 ($n = 6195$); Ω_2 corresponds to the period January 4th, 2011 to December 12th, 2012 ($n = 7189$); Ω_0^{stress} corresponds to the period $\Omega_0 \cup$ {January 24th, 2007 to September 30th, 2007 ($n = 4851$)}; $\Omega_0'^{stress}$ corresponds to the period $\Omega_0 \cup$ { January 24th, 2007 to December 31th, 2007 ($n = 5451$)}.

4 Conclusion

In this paper, we have studied the asymptotic Gaussian property given by the Theorem 2.1 for the distribution of the estimator $\widehat{VaR}_p$ of VaR_p, which allows to build a confidence interval CI_q around VaR_p. Since the convergence speed of this property depends on the unknown distribution f , the sample size n of the data set and the confidence level p of VaR_p we have provided a comprehensive analysis of the feasibility of the Theorem 2.1 for finite samples with a panel of distributions on a data set of operational risks. This first work shows the limit of the theoretical results as soon as n is small (less than 1000) and the distributions characterised by fat-tail.

In a second part, we use Theorem 2.1 to build a confidence interval CI_q around VaR_p and we use this confidence interval to determine an area whose lower bound is VaR_p and upper bound the upper bound of the confidence interval: this area corresponds to a new alert indicator SSVaR that we can use to detect anormal events in a very quick way using a dynamical procedure as described in section 3 (using a data set composed by operational risks).

We have observed that the choice of the distributions which characterise the risk factors were crucial in the building of the SSVaR, and also when we use it in a dynamical way. This last approach permits to introduce a new methodology for stress testing.

References

[1] Bank for international settlements. Basel III: A global regulatory framework for more resilient banks and banking systems. *Basel Committee on Banking Supervision*, 2011, Basel, Switzerland.

[2] Ergashev B. Abdymomunov A., Blei S. Integrating stress scenarios into risk quantification models. *Journal of Financial Services Research*, 47(1), 57-79, 2015.

[3] Guégan D. Hassani B. Stress testing engineering: the real risk measurement. *Future Perspective in Risk Models and Finance eds A. Bensoussan, D. Guégan, C. Tapiero*, 2015, New York, USA.

[4] Ledermann D. Alexander C. ROM simulation: applications to stress testing and VaR. *ICMA Centre Discussion Papers in Finance*, icma-dp2012, (09), 2012, Reading, UK.

[5] Loader C. R. Bandwidth selection: classical or plug-in? *The Annals of Statistics*, 27(2), 415-438, 1999.

[6] McCulloch J. H. Simple consistent estimators of stable distribution parameters. *Communications in Statistics*, Simulations 15, 1109–1136, 1986.

[7] Wong C. Fong T. P. Stress testing banks' credit risk using mixture vector autoregressive models. *Hong Kong monetary authority working paper*, 2008(13), Hong Kong, China.

[8] Rao C. R. Linear statistical inference and its applications. *Wiley-Interscience*, 2001, New York, USA.

[9] Sheather S. J. Density estimation. *Statistical Science*, 19(4), 588–597, 2004.

[10] Silverman B. W. Density estimation for statistics and data analysis, *Chapman and Hall*, 1986, London, UK.

[11] Zhao X. Gao F. Delta method in large deviations aand moderate deviations for estimations. *The Annals of Statistics*, 39, 2, 1211–1240, 2011.

[12] Zicchino L. Hoggarth G., Sorensen S. Stress tests of UK banks using a var approach. *Bank of England's working paper*, 2005.

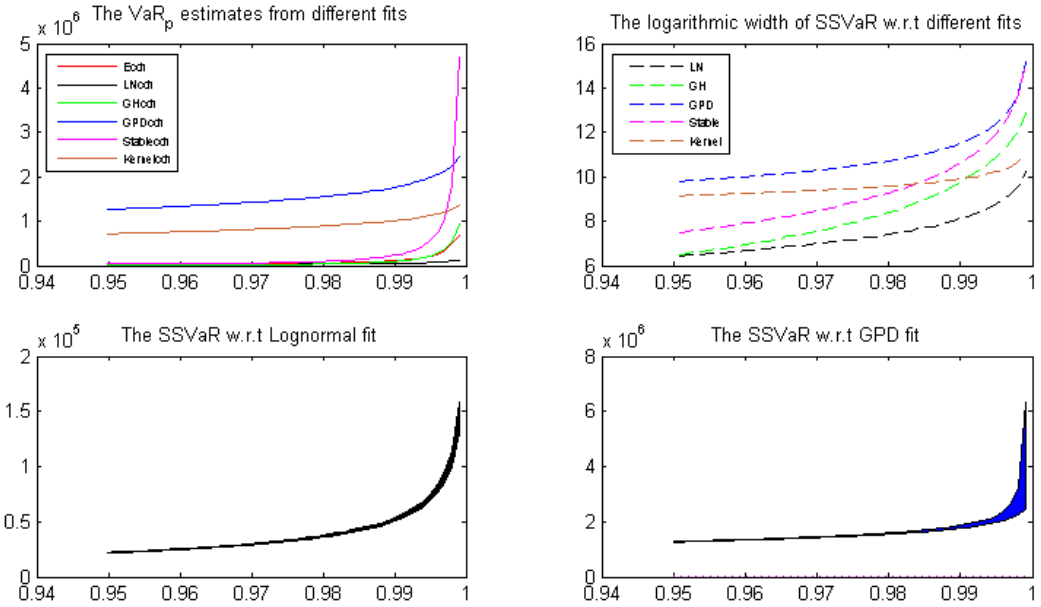


Figure 3: Top left corresponds to the VaR_p for six distributions where p varies from 0.95 to 0.999. For lognormal we use black; for GH we use green; for GPD we use blue; for $\alpha - stable$ we use purple; for Kernel we use brown; for empirical we use red. Top right corresponds to the width of the SSVaR for the same distributions. Bottom left corresponds to the SSVaR when we fit on Ω_2 a lognormal distribution. Bottom right corresponds to the SSVaR when we fit on Ω_2 a GPD distribution.

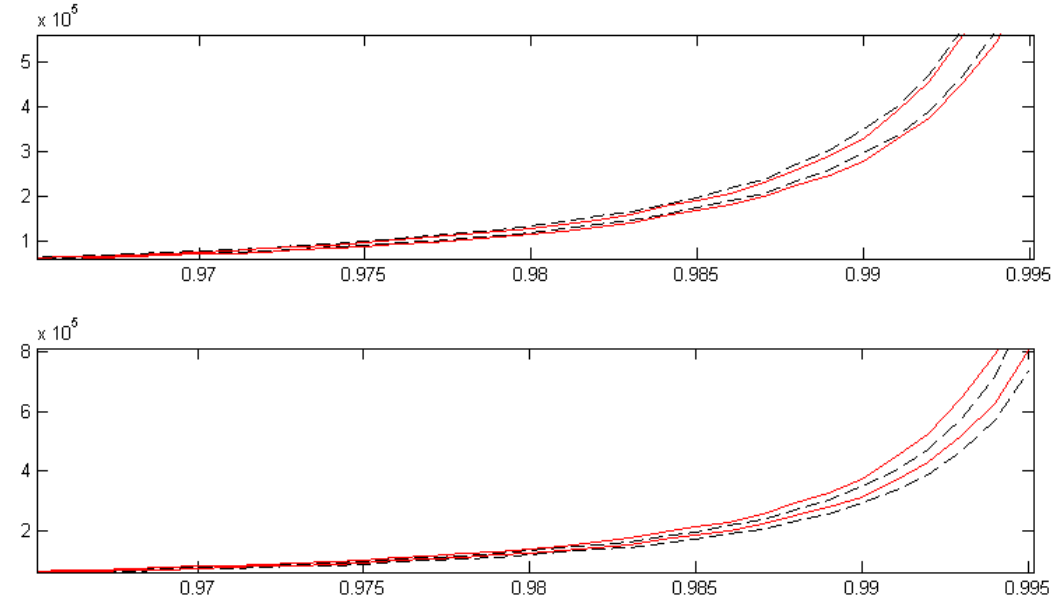


Figure 4: Top: we represent the SSVaR; dot lines correspond to the SSVaR built with Ω_0 on which an $\alpha - stable$ distribution is fitted, and solid lines correspond to the SSVaR built with Ω_1 still using an $\alpha - stable$ distribution. Bottom: we represent SSVaR, dot lines using Ω_0 and solid lines using Ω_0^{stress}, still fitting an $\alpha - stable$ distribution.

	μ	σ			
Lognormal	7.7853	1.6100			
	(0.0289)	(0.0204)			
	α	β	δ	μ	λ
GH	8.7572e-4	8.7426e-4	1.1066e+3	9.6355e+2	-7.1308e-1
	(1.2794e-4)	(1.2814e-4)	(1.7980e+2)	(1.4645e+2)	(7.7093e-2)
	ξ	β	μ		
GPD	8.9810e-1	3.7905e+4	44000		
	(1.2786e-1)	(4.8380e+3)			
	ξ	β	μ	block length	
GEV	8.8767e-1	2.2458e+3	2.0498e+3	4	
	(3.8743e-2)	(1.1887e+1)	(8.8251e+1)		
	α	β	γ	δ	
$\alpha - stable$	0.800	0.792	1085.317	1677.545	
	(0.0223)	(0.0252)	(30.8179)	(99.3442)	

Table 3: Values of the estimated parameters for Ω_1. We use maximum likelihood estimation to estimate the Lognormal, GH, GPD (except the location) and GEV parameters. The location of the GPD is estimated using Hill plot. For estimating the parameters of the $\alpha - stable$ distribution, we use McCulloch method. The standard deviation of the Lognormal and GEV parameter estimates are computed through the inverse of the Hessian matrix. For GH, GPD and $\alpha - stable$, they are computed using Bootstrap algorithm.

	μ	σ			
Lognormal	7.9491	1.2373			
	0.0068	0.0048			
	α	β	δ	μ	λ
GH	2.8367e-5	2.7881e-5	1000	1000	-0.5
	(1.3616e-6)	(6.0339e-8)	(59.2976)	(0)	(0.0756)
	ξ	β	μ		
GPD	-1.0080e-2	3.2245e+5	314391.3		
	(0.0093)	(2.8423e+4)			
	ξ	β	μ	block length	
GEV	0.9853	2635,2984	2583.3371	9	
	(0.0166)	(64.3526)	(46.4861)		
	α	β	γ	δ	
$\alpha - stable$	0.7720	0.95	908.5263	1655.948	
	(0.0056)	(0)	(9.1110)	(18.5196)	

Table 4: Values of the estimated parameters for Ω_2. We use maximum likelihood estimation to estimate the Lognormal, GH, GPD (except the location) and GEV parameters. The location of the GPD is estimated using Hill plot. For estimating the parameters of the $\alpha - stable$ distribution, we use McCulloch method. The standard deviation of the Lognormal and GEV parameter estimates are computed through the inverse of the Hessian matrix. For GH, GPD and $\alpha - stable$, they are computed using Bootstrap algorithm.

	p=0.95	p=0.975	p=0.99	p=0.995	p=0.999
Lognormal	3.4021e+4 [2.9947e+4,3.8095e+4] (0.041)	5.6762e+4 [4.6885e+4,6.6638e+4] (0.022)	1.0120e+5 [7.3343e+4,1.2905e+5] (0.010)	1.5302e+5 [9.1873e+4,2.1418e+5] (0.005)	3.3868e+5 [5.1833e+4,6.2552e+5] (0.008)
GH	2.4535e+4 [2.1064e+4,2.8005e+4] (0.087)	5.1868e+4 [4.0925e+4,6.2812e+4] (0.073)	1.2168e+5 [8.2014e+4,1.6134e+5] (0.030)	2.1548e+5 [1.2595e+5,3.0500e+5] (0.033)	5.4143e+5 [1.6253e+5,9.2032e+5] (0.021)
GPD	6.2543e+5 [5.3937e+5,7.1149e+5] (0.049)	1.1715e+6 [9.0531e+5,1.4377e+6] (0.021)	2.6455e+6 [1.5527e+6,3.7383e+6] (0.006)	5.0525e+6 [1.7319e+6,8.3731e+6] (0.012)	NA
GEV	3.2211e+4 [2.7702e+4,3.6719e+4] (0.042)	6.0755e+4 [4.6921e+4,7.4588e+4] (0.028)	1.3736e+5 [8.0850e+4,1.9387e+5] (0.015)	2.5892e+5 [9.1556e+4,4.2629e+5] (0.018)	NA
$\alpha - stable$	3.1485e+4 [2.6338e+4,3.6632e+4] (0.046)	6.8617e+4 [4.8567e+4,8.8666e+4] [0.029]	2.0340e+5 [8.9014e+4,3.1780e+5] (0.023)	4.9410e+5 [4.2248e+4,9.4595e+5] (0.016)	NA
Kernel	5.7488e+4 [5.4316e+4,6.0660e+4] (0.045)	7.4460e+4 [6.6747e+4,8.2173e+4] (0.024)	2.4289e+5 [2.1320e+5,2.7257e+5] (0.019)	7.2827e+5 [6.4861e+5,8.0793e+5] (0.006)	9.0110e+5 [6.9982e+5,1.1024e+6] (0.002)

Table 5: Building of CI_q using six distributions fitted on Ω_1 ($n = 3109$) for five values of p. In each cell, we provide on the first line $\bar{\Psi}$, on the second line CI_q and on the third line, in bracket, $\hat{q}$.

	p=0.95	p=0.975	p=0.99	p=0.995	p=0.999
Lognormal	2.1686e+4 [2.1074e+4,2.2301e+4] (0.040)	3.2058e+4 [3.0743e+4,3.3372e+4] (0.024)	5.0416e+4 [4.7098e+4,5.3734e+4] (0.004)	6.8612e+4 [6.2177e+4,7.5044e+4] (0.011)	1.3055e+5 [1.0245e+5,1.5864e+5] (0.003)
GH	1.1257e+4 [1.0621e+4,1.1894e+4] (0.055)	2.7273e+4 [2.4463e+4,3.0084e+4] (0.011)	9.2949e+4 [7.5722e+4,1.1018e+5] (0.009)	2.1762e+5 [1.6323e+5,2.7199e+5] (0.003)	9.5403e+5 [5.8866e+5,1.3149e+6] (0.007)
GPD	1.1482e+6 [1.1370e+6,1.1595e+6] (0.037)	1.3076e+6 [1.2904e+6,1.3248e+6] (0.025)	1.5025e+6 [1.4738e+6,1.5312e+6] (0.010)	1.6386e+6 [1.5970e+6,1.6801e+6] (0.001)	1.9191e+6 [1.8265e+6,2.0117e+6] (0)
GEV	4.9012e+4 [4.6670e+4,5.1355e+4] (0.038)	9.8218e+4 [9.0633e+4,1.0580e+5] (0.017)	2.4382e+5 [2.0963e+5,2.7801e+5] (0.009)	4.8468e+5 [3.7961e+5,5.8975e+5] (0.008)	2.4311e+6 [1.0139e+6,3.8483e+6] (0.006)
α_{stable}	3.2957e+4 [3.1217e+4,3.4698e+4] (0.045)	7.4012e+4 [6.7133e+4,8.0912e+4] (0.020)	2.2930e+5 [1.8839e+5,2.7020e+5] (0.010)	5.5342e+5 [3.9828e+5,7.0857e+5] (0.011)	4.5567e+6 [1.0639e+6,8.0494e+6] (0.009)
Kernel	4.7512e+4 [4.6684e+4,4.8340e+4] (0.043)	1.0166e+5 [9.9397e+4,1.0393e+5] (0.027)	2.7763e+5 [2.6919e+5,2.8607e+5] (0.009)	6.8311e+5 [6.6027e+5,7.0595e+5] (0.008)	1.9261e+6 (1.8142e+6,2.0380e+6) (0.001)

Table 6: Building of CI_q using six distributions fitted on Ω_2 ($n = 32727$) for five values of p. In each cell, we provide on the first line $\bar{\Psi}$, on the second line CI_q and on the third line, in bracket, $\hat{q}$.

Appendices

.A Proof of Theorem 1

It is based on part of the works of Rao (2001) [8]. Here we provide a new version proof with more details. Consider the r.v. $\hat{\xi}_{p_n}$,

$$P(\hat{\xi}_{p_n} \leq x) = P(x_{[np]+1} \leq x) \tag{.1}$$

And by the definition of the order statistics, $x_{[np]+1} \leq x \Longleftrightarrow \sharp(observations \leq x) \geq [np]+1$

$$P(\sharp(observations \leq x) \geq [np]+1) = \sum_{r=m}^{n} C_n^r (F(x))^r (1-F(x))^{n-r} \tag{.2}$$

where $m = [np]+1$, since the proof focusing on the case $n \to \infty$, for simplify without loss of generality, we assume that $m = np$. Then, using the beta function and incomplete beta function:

$$\begin{aligned} F_{\hat{\xi}_{p_n}}(x) =& P(\hat{\xi}_{p_n} \leq x) = \sum_{r=m}^{n} C_n^r (F(x))^r (1-F(x))^{n-r} \\ =& I_{F(x)}(m, n-m+1) = \frac{B_{F(x)}(m, n-m+1)}{B(m, n-m+1)} \\ =& \frac{n!}{(m-1)!(n-m)!} \int_0^{F(x)} t^{m-1}(1-t)^{n-m} dt \end{aligned} \tag{.3}$$

To get the density function of $\hat{\xi}_{p_n}$, differentiate $F_{\hat{\xi}_{p_n}}(x)$ w.r.t x,

$$f_{\hat{\xi}_{p_n}}(x) = \frac{n!}{(m-1)!(n-m)!} F(x)^{m-1}(1-F(x))^{n-m} f(x) \tag{.4}$$

Let $H = \hat{\xi}_{p_n}$, $Y = F_X(\hat{\xi}_{p_n}) = F_X(H)$:

$$F_Y(y) = P(Y \leq y) = P(F_X(\hat{\xi}_{p_n}) \leq y) = P(\hat{\xi}_{p_n} \leq F_X^{-1}(y)) = F_H(F_X^{-1}(y)) \tag{.5}$$

It's equivalent to:

$$F_Y(y) = F_H(F_X^{-1}(y)) \tag{.6}$$

So the density function of Y is:

$$\frac{dF_Y(y)}{dy} = \frac{dF_H(F_X^{-1}(y))}{dy} = \frac{dF_H(F_X^{-1}(y))}{dF_X^{-1}(y)} \cdot \frac{dF_X^{-1}(y)}{dy} \tag{.7}$$

and $\frac{dF_X^{-1}(y)}{dy} = \frac{1}{f(F_X^{-1}(y))}$, also from equation (.7),

$$\begin{aligned} f_Y(y) = \frac{dF_Y(y)}{dy} &= f_H(F_X^{-1}(y)) \cdot \frac{1}{f(F_X^{-1}(y))} = f_{\hat{\xi}_{p_n}}(F_X^{-1}(y)) \cdot \frac{1}{f(F_X^{-1}(y))} \\ &= \frac{n!}{(m-1)!(n-m)!} y^{m-1}(1-y)^{n-m} \end{aligned} \tag{.8}$$

Now, introduce a new r.v. $Z = \frac{\sqrt{n}(Y-p)}{\sqrt{pq}}$, then the density function of Z is:

$$A \times B \tag{.9}$$

where

$$\begin{aligned} A &= \frac{1}{\sqrt{n}} \frac{n!}{(m-1)!(n-m)!} p^{\frac{m-1}{2}} q^{\frac{n-m+1}{2}} \\ B &= (1 + \frac{z\sqrt{q}}{\sqrt{np}})^{m-1} (1 - \frac{z\sqrt{p}}{\sqrt{nq}})^{n-m} \end{aligned} \tag{.10}$$

Take the logarithm of (.9),

- The term $ln(A)$ which is not involving z is:

$$ln(A) \xrightarrow{n\to\infty} Constant \tag{.11}$$

- The term $ln(B)$, using Taylor expansion of ln function:

$$\begin{aligned} ln(B) &= (m-1)ln(1 + \frac{z\sqrt{q}}{\sqrt{np}}) + (n-m)ln(1 - \frac{z\sqrt{p}}{\sqrt{nq}}) \\ &= (np-1)ln(1 + \frac{z\sqrt{q}}{\sqrt{np}}) + (nq)ln(1 - \frac{z\sqrt{p}}{\sqrt{nq}}) \\ &= (np-1)(\frac{z\sqrt{q}}{\sqrt{np}} - \frac{z^2 q}{2np}) + (nq)(-\frac{z\sqrt{p}}{\sqrt{nq}} - \frac{z^2 p}{2nq}) \\ &= z\sqrt{npq} - \frac{z\sqrt{q}}{\sqrt{np}} - \frac{1}{2}z^2 q + \frac{z^2 q}{2np} - z\sqrt{npq} - \frac{1}{2}z^2 p \\ &= -\frac{1}{2}z^2 - \frac{z\sqrt{q}}{\sqrt{np}} + \frac{z^2 q}{2np} \xrightarrow{n\to\infty} -\frac{1}{2}z^2 \end{aligned} \tag{.12}$$

So the density function of Z tends to $Constant \times e^{-\frac{1}{2}z^2}$ as $n \to \infty$. Here, remind that the characteristic function of $\mathcal{N}(0,1)$ is $e^{-\frac{t^2}{2}}$, which dose NOT depend on the constant $\frac{1}{\sqrt{2\pi}}$. Hence by Scheffe's theorem(Rao (2001) [8], (xv) 2c.4), the asymptotic distribution is $\mathcal{N}(0,1)$. Remind that $y = p(1 + z\sqrt{\frac{q}{np}})$. Hence the a.d. of $\sqrt{n}(y-p)$ is $\mathcal{N}(0, pq)$.
Remind that:

$$\begin{aligned} &Y = F_X(\hat{\xi}_{p_n}) \quad \Longrightarrow \quad \hat{\xi}_{p_n} = F_X^{-1}(Y) \\ &\xi_p = F_X^{-1}(p) \end{aligned} \tag{.13}$$

then apply Rao (2001) [8] (i) 6a.2, we have:

$$\sqrt{n}(\hat{\xi}_{p_n} - \xi_p) \xrightarrow{distribution} \mathcal{N}\left(0, pq((F_X^{-1}(p))\prime)^2\right) \tag{.14}$$

since $(F_X^{-1}(p))\prime = \frac{1}{f(F_X^{-1}(p))} = \frac{1}{f(\xi_p)}$
Finally,

$$\hat{\xi}_{p_n} \xrightarrow{distribution} \mathcal{N}\left(\xi_p, \frac{p(1-p)}{f(\xi_p)^2 n}\right) \qquad \blacksquare \tag{.15}$$

.B Methodology for non parametric fitting

Given $\Omega_i = \{X_1^i, ..., X_n^i\}, i = 1, 2$ sequences of i.i.d r.v. from f. Now we describe the methodology of non-parametric fit as following:

1. The Kernel density estimate of f at the point x_0 is given by:

$$\hat{f}_h(x_0) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x_0 - X_i}{h}\right) \tag{.16}$$

where K is the Kernel function and h is the bandwidth. Sheather S. J. (2004) argues that the value of the bandwidth is of critical importance, while the shape of the Kernel function has little practical impact. But we observe that in practice when we analyse the data set with "extreme points", the choice of Kernel function is also crucial. Since the original idea of Kernel function is from the definition of a general weight function which is symmetric w.r.t. 0 , $\int_{support} K\left(\frac{x-x_0}{h}\right) dx = 1$ and $\int_{support} xK\left(\frac{x-x_0}{h}\right) dx = 0$, to make sure that most of the weight of the density estimation at point x_0 falls near x_0 (Silverman B. W. (1986)). Before defining a new Kernel function to make the $\hat{f}_h$ adapt to the extreme points area, we discuss about the bias of $\hat{f}_h$ first. Define $\int t^2 K(t)dt = k_2$, where $K(.)$ is the Kernel function. Thus k_2 could be explained as the variance regarding $K(.)$. Then the bias in the estimation of f at x_0 is:

$$bias(x_0) = \left|E(\hat{f}_h) - f(x)\right| = \left|\int \frac{1}{h} K\left(\frac{x_0 - x}{h}\right) f(x)dx - f(x_0)\right| \tag{.17}$$

let $x = x_0 - ht$, then:

$$bias(x_0) = \left|-\int K(t) f(x_0 - ht)dt - f(x_0)\right| \tag{.18}$$

From Taylor expansion we have $f(x_0 - ht) = f(x_0) - htf'(x) + \frac{1}{2}h^2t^2f''(x) + ...$, then:

$$\begin{aligned} bias(x_0) &\approx \left|-f(x_0)\int K(t)dt + hf'(x_0)\int tK(t)dt - \frac{1}{2}h^2 f''(x_0)\int t^2K(t)dt - f(x_0)\right| \\ &= \left|-2f(x_0) - \frac{1}{2}h^2 f''(x_0)k_2\right| \end{aligned} \tag{.19}$$

Thus the bias is positively related to h^2 and k_2 which could be a trade off. Since h is with power 2, we may give the priority to have a low h even if this decision leads to a relatively high k_2. Thus we define the so called "scaled Gaussian Kernel" (Gaussian density with location 0 and scaled by the sample standard deviation.) to estimate the Kernel density at the extreme points.

2. Based on the minimization of the asymptotic mean integrated squared error (AMISE) and the rules of thumb, we get the bandwidth $h_0 = 0.9An^{-0.2}$, where

$$A = min(S, (sample\, interquantile\, range)/1.34) \tag{.20}$$

under the assumption that f is Gaussian (Silverman (1986) [10]). It could be a benchmark (or initial value) for h.

3. To adapt the bandwidth to a specific data set, we use the cross validation rule given by Loader (1999) [5] and take the log:

$$log(CV(h)) = \sum_{i=1}^{n} log\left(\hat{f}_{h,-i}(X_i)\right) \tag{.21}$$

where $\hat{f}_{h,-i}(X_i)$ denotes the Kernel density estimate with data point X_i deleted. Beginning with h_0, by finding the local log maximal of CV, we get the local optimal bandwidth h_{opt}. We only consider the local maximal since the bandwidth can neither too large (leads to large bias) nor too small (leads to large variance).

4. After choosing the Kernel function and the bin, by equation (5), we have the Kernel estimator of f at x_0.

.C Bootstrapped methodology for random generation from Gaussian Kernel fit

For Gaussian Kernel, the random generation is replaced by a bootstrapped methodology in given by Silverman (1986) [10] as following: given p, n and a data set Ω with i.i.d sample,

1. choose X_i $(i = 1, ..., n * 1000)$ uniformly with replacement from Ω.
2. generate ϵ_i $(i = 1, ..., n * 1000)$ from Gaussian Kernel.
3. define $\bar{X} = mean(\Omega)$, then the new sample set with size n*1000 is given below:

$$\bar{X} + \left(X_i - \bar{X} + h\epsilon_i\right) / \left(1 + h^2\right)^{\frac{1}{2}}, i = 1, ..., n * 1000 \tag{.22}$$

4. we put this new sample set into a n*1000 matrix and sort it in ascending order. According to the definition of $\widehat{VaR}_p$, we have its 1000 i.i.d realizations which is the $[np] + 1$ row of this matrix.